

Sistemas Inteligentes. Aplicaciones en Minería de Datos, Robótica Evolutiva y Redes de Computadoras

**Laura Lanzarini¹, Waldo Hasperue², Hernán Vinuesa³, Leonardo Corbalán⁴, Germán Osella Massa⁵,
{laural, whasperue, hvinuesa, corbalan, gosella}@lidi.info.unlp.edu.ar**

**Instituto de Investigación en Informática LIDI (III-LIDI)⁶.
Facultad de Informática. UNLP**

CONTEXTO

Esta presentación corresponde al Subproyecto “Sistemas Inteligentes” perteneciente al Proyecto “Algoritmos Distribuidos y Paralelos. Aplicación a Sistemas Inteligentes y Tratamiento Masivo de Datos” del Instituto de Investigación en Informática LIDI.

RESUMEN

Esta línea de investigación se centra en el estudio y desarrollo de Sistemas Inteligentes basados en mecanismos de adaptación. Actualmente el énfasis está puesto en la transferencia de tecnología a las áreas de minería de datos, robótica evolutiva y redes de computadoras.

En el área de la Minería de Datos, los temas centrales se encuentran relacionados con la investigación de nuevas estrategias adaptativas que definan las acciones a seguir para lograr el beneficio esperado a partir de un modelo de la información basado en reglas de clasificación, facilitando de esta manera su interpretación.

En el área de la robótica evolutiva, el énfasis está puesto en la construcción de un controlador con capacidad de adaptación a entornos dinámicos. Se trabaja en el desarrollo de nuevos métodos para la resolución de problemas utilizando agentes capaces de percibir y actuar en entornos complejos cuyos resultados son aplicados directamente en esta área.

La aplicación de distintas metaheurísticas sobre redes Peer-to-Peer (P2P), es otra línea de investigación que se está llevando a cabo en el III-LIDI. El objetivo central es mejorar la capacidad de adecuación del sistema a los cambios rápidos del entorno de información.

Palabras claves : Redes Neuronales, Algoritmos Evolutivos, Minería de Datos, Optimización mediante Cúmulos de Partículas, Colonias de Hormigas.

1. INTRODUCCIÓN

Los Sistemas Inteligentes han demostrado ser herramientas sumamente útiles en la resolución de problemas complejos. Su capacidad de adaptación al entorno de información les permite brindar buenos resultados en distintas áreas.

En el Instituto de Investigación en Informática LIDI se está trabajando en este tema desde hace varios años. Inicialmente se desarrollaron estrategias basadas en Redes Neuronales y Algoritmos Evolutivos aplicables al Reconocimiento de Patrones [Lan00, Lan04] así como al control de agentes autónomos [Cor03a, Cor03b, Oli05]. Luego se profundizó en el estudio de la capacidad de caracterización de este tipo de estrategias a fin de establecer un modelo de la información disponible. Se definieron y desarrollaron varias estrategias basadas en redes neuronales competitivas con aplicaciones concretas en distintas áreas [Has07a, Cor06a,

¹ Profesor Titular. Facultad de Informática. UNLP

² Becario de Estudio de la CIC - Facultad de Informática. UNLP

³ Becario de Entrenamiento de la CIC - Facultad de Informática. UNLP

⁴ Becario de Perfeccionamiento de la UNLP. Jefe de Trabajos Prácticos - Facultad de Informática. UNLP

⁵ Becario Doctoral de CONICET. Jefe de Trabajos Prácticos - Facultad de Informática. UNLP

⁶ Calle 50 y 115 1er Piso, (1900) La Plata, Argentina, TE/Fax +(54) (221) 422-7707. <http://weblidi.info.unlp.edu.ar>

Cor06b]. Actualmente, con el objetivo de mejorar la transferencia tecnológica de los resultados obtenidos, el énfasis está puesto por un lado, en la definición de estrategias que faciliten la interpretación del modelo y por otro en la incorporación de mecanismos de adaptación a entornos dinámicos.

A continuación se detallan brevemente los avances realizados últimamente.

1.1. Minería de Datos

Las investigaciones realizadas en el III-LIDI relacionadas con diferentes mecanismos de aprendizaje y adaptación de redes neuronales competitivas aplicables a la Minería de Datos han facilitado la construcción de modelos de la información disponible principalmente a partir de reglas de asociación y clasificación [Has05, Has06]. Sin embargo, existen situaciones en las cuales el modelo por si solo es incapaz de transmitir al usuario el conocimiento adquirido así como su importancia y relación con los datos subyacentes. En esta dirección, las técnicas visuales resultan una manera intuitiva de presentación, dándole una perspectiva exploratoria y contextual.

El gran obstáculo que actualmente presenta la transferencia tecnológica en temas referidos a estrategias de minería de datos, es el desconocimiento por parte del usuario acerca de la forma de adaptar el modelo a sus necesidades. Es por eso que resulta de fundamental importancia incentivar el estudio de estrategias que permitan obtener, a partir de la información del beneficio esperado, las acciones necesarias para modificar los datos y por lo tanto, lograr cambiar el modelo adecuadamente.

Como forma de ayudar a revertir esta situación, se ha desarrollado una estrategia que define, a partir del modelo basado en reglas de clasificación, las acciones a seguir para lograr el beneficio esperado. Para construir las reglas se utilizó una matriz de co-asociación cuyos valores surgen de la combinación de distintos métodos de clustering aplicados a los datos de entrada.

De esta manera, aunque la decisión final continua estableciéndola el usuario, puede disponerse de una línea de acción [Has07b].

1.2. Robótica Evolutiva

En el III-LIDI se está trabajando desde el año 2002 en el desarrollo de controladores para agentes autónomos utilizando Redes Neuronales y Algoritmos Evolutivos. Las investigaciones realizadas han demostrado que la descomposición del problema en tareas más simples, facilita el desarrollo del controlador. Los métodos propuestos inicialmente estuvieron basados en aprendizaje incremental [Cor03b, Cor05] y aprendizaje por capas [Oli05]. Ambos requieren una gran asistencia, en el primer caso para identificar las distintas etapas del aprendizaje - no siempre factibles de ser establecidas - y en el segundo caso para integrar las capas cuyo comportamiento fue aprendido en un orden específico. Como solución a esto, se desarrolló un mecanismo de integración automática de módulos que permitió combinar comportamientos básicos aprendidos previamente [Ose06] reduciendo de esta forma el costo de entrenamiento.

Analizando globalmente las soluciones mencionadas anteriormente, puede observarse que concentran la aplicación de estrategias adaptativas en la fase de generación del controlador no permitiendo realizar adaptaciones con posterioridad. Esto perjudica la aplicación del controlador en ambientes dinámicos.

Actualmente se está trabajando en la evolución del controlador a lo largo de su vida útil combinando un método basado en evolución de módulos neuronales con un algoritmo evolutivo específico. El primer método es utilizado para producir el controlador mientras que

el segundo lo ajusta durante su funcionamiento. Como resultado, se obtiene un controlador basado en una red neuronal que posee una arquitectura mínima y capacidad de adaptación en la fase de ejecución [Vin07a]. También se han analizado distintas estrategias elitistas de evolución como forma de mejorar el efecto de los operadores genéticos [Vin07b].

1.3. Redes de Computadoras

La aplicación de distintas metaheurísticas sobre redes Peer-to-Peer (P2P), es otra línea de investigación que se está llevando a cabo en el III-LIDI. Las redes P2P constituyen un sistema distribuido, descentralizado y sumamente dinámico, donde la información disponible cambia con mucha rapidez, al igual que la topología misma de la red. La recuperación de información y en general la búsqueda de recursos en estos sistemas totalmente descentralizados es compleja.

El objetivo central de esta investigación es mejorar la capacidad de adecuación del sistema a los cambios rápidos del entorno de información utilizando estrategias basadas en Redes Neuronales, Algoritmos Evolutivos, Colonias de Hormigas (ACO - Ant Colony Optimization) y Optimización Basada en Cúmulos de Partículas (PSO - Particle Swarm Optimization).

Actualmente se está trabajando en mejorar el algoritmo de aprendizaje, por el cual se mantienen actualizados los pesos de las conexiones en cada nodo, con alguna estrategia ACO y el manejo de vecindad con una nueva variante basada en PSO.

Sin embargo, los sistemas PSO pueden resultar apropiados si se cuenta con la posibilidad de realizar algún análisis previo a la incorporación de un nodo a la red P2P. En [Shi07] se ha utilizado PSO para determinar la interconexión de los nodos en este tipo de redes con algunas restricciones. Una vez que el nodo es incorporado, formando parte del sistema P2P, una red neuronal de tipo LVQ es utilizada para dirigir las búsquedas sólo hacia los nodos vecinos más prometedores, obteniéndose así una reducción drástica en el tráfico generado para estos fines. Esto redundaría en un mayor rendimiento del sistema P2P en general.

Investigaciones recientes llevadas a cabo en el III-LIDI han permitido definir y desarrollar una extensión original de PSO que, a diferencia de [Ken95][Shi99], permite trabajar con una población de tamaño variable. De esta forma, no es necesario definir a priori la cantidad de soluciones a utilizar, evitando así condicionar la calidad de la solución a obtener [Lez07]. Esta estrategia ha sido probada con buenos resultados en el área de optimización de funciones y se espera aplicarla con éxito en el proceso de construcción de la topología del sistema P2P.

2. TEMAS DE INVESTIGACIÓN Y DESARROLLO

- Análisis de los resultados obtenidos de la combinación de distintas arquitecturas para resolver un único problema. Interesa especialmente la reducción del tiempo de adaptación.
- Estudio e investigación de experiencias realizadas en la obtención de controladores en el área de la robótica evolutiva haciendo hincapié en el dinamismo de la representación y en el conocimiento previo necesario para resolver el problema.
- Investigación de nuevas estrategias evolutivas para resolver problemas complejos a través de la combinación de controladores elementales buscando reusar el conocimiento adquirido minimizando el tiempo de adaptación.
- Estudio y análisis de diferentes estructuras de modelización dinámicas. Interesa especialmente la obtención de árboles de decisión incluyendo los árboles difusos e incrementales.

- Desarrollo y aplicación de diferentes de métricas que permitan analizar el conjunto de reglas a utilizar para representar el modelo. Esto incluye considerar distintas estrategias de poda que permitan maximizar la representación obtenida.
- Estudio, análisis y especificación del concepto de “acción” a partir de los datos modelizados buscando mantener independencia del dominio de aplicación.
- Análisis y especificación de las características accionables. Análisis de casos con y sin restricciones. Efectos sobre el modelo.
- Estudio e investigación de estrategias para la obtención de patrones de acción a partir del modelo, de una identificación adecuada de las características accionables de los datos y del beneficio a maximizar.
- Estudio de distintas metaheurísticas de optimización basadas en trayectoria y en población aplicables al problema de ruteo en redes P2P.
- Análisis de la importancia de los parámetros de las distintas metaheurísticas en el funcionamiento y eficiencia de la estrategia seleccionada. Análisis de la función de aptitud a utilizar en el caso de redes P2P.
- Estudios de performance de los algoritmos desarrollados. Análisis de eficiencia en la resolución de problemas concretos.

3. RESULTADOS OBTENIDOS/ ESPERADOS.

- Definición e implementación de una estrategia de evolución elitista aplicable a la adaptación de redes neuronales. En este caso, resulta de interés la reducción de los efectos negativos de los operadores genéticos durante el proceso de adaptación.
- Desarrollo e implementación de una estrategia de integración de los módulos evolutivos. Interesa especialmente mantener el dinamismo de la arquitectura a fin de no limitar sus capacidades.
- Desarrollo e implementación de una estrategia para la obtención de reglas de clasificación a partir de una matriz de co-asociación.
- Desarrollo e implementación de mecanismos que permitan la especificación de un conjunto de acciones a seguir a fin de objetivar la interpretación de la información modelizada.
- Análisis y ampliación de los frameworks existentes para Robótica Evolutiva en lo que hace a la definición de escenarios, interacción con robots específicos, plataformas de desarrollo y posibilidades de desarrollos multi-agentes.
- Desarrollo e implementación de una estrategia basada en cúmulos de partículas (PSO) con tamaño de población variable basado en los conceptos de edad y vecindario. Se ha comprobado que el mecanismo utilizado para incorporar nuevos individuos así como la forma de calcular el tiempo de vida preserva la diversidad de la población.
- Resolución de problemas concretos, tanto en ambientes simulados como en el mundo real. En este último caso, resulta de fundamental importancia la optimización del algoritmo propuesto.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro de los temas involucrados en esta línea de investigación se están desarrollando actualmente 2 tesis de doctorado, 1 una de maestría y al menos 2 tesinas de grado de Licenciatura.

También participan en el desarrollo de las tareas becarios y pasantes del III-LIDI.

5. REFERENCIAS

- [Cor03a] Corbalán, Lanzarini. GNE. Grupos Neuronales Evolutivos. *XXIX Conferencia Latinoamericana de Informática*. CLEI 2003. Bolivia. 2003
- [Cor03b] Corbalán, Lanzarini. Evolving Neural Arrays A new mechanism for learning complex action sequences. *Special Issue of Best Papers presented at CLEI'2002*. Volumen 6. Number 1, December 2003. Montevideo, Uruguay.
- [Cor05] Corbalán, Lanzarini, De Giusti A. ALENA. Adaptive-Legth Evolving Neural Arrays". *Journal of Computer Science and Technology*. Vol 5, nro. 4. 2005. Pags. 59-65.
- [Cor06a] Corbalán, Osella Massa, Russo, Lanzarini. Image recovery using a new nonlinear Adaptive Filter based on Neural Networks. *CIT - Journal of Computing and Information Technology*. Vol. 14, No.4. December 2006 - Pág. 315-320.
- [Cor06b] Corbalán, Hasperue, Osella Massa, Lanzarini. BPNn-CPN. Nuevo método para segmentación de Imágenes basado en Redes Neuronales Artificiales. *IV Workshop de Computación Gráfica, Imágenes y Visualización (WCGIV)*. CACIC 2006. San Luís. Argentina. Octubre de 2006.
- [Has05] Hasperué, Lanzarini. Dynamic Self-Organizing Maps. A new strategy to enhance topology preservation. *XXXI Conferencia Latinoamericana de Informática*. CLEI 2005.
- [Has06] Hasperué, Lanzarini. Classification Rules obtained from Dynamic Self-organizing Maps. *VII Workshop de Agentes y Sistemas Inteligentes*. CACIC 2006. San Luis. Argentina. Octubre de 2006.
- [Has07a] Hasperue, Corbalán, Lanzarini, Bría. Skeletonization of Sparse Shapes using Dynamic Competitive Neural Networks. *Ibero-American Journal of Artificial Intelligence*. Vol.11. Nro.35. pp.33-42 2007.
- [Has07b] Hasperué, Lanzarini. Extracting Actions from Classification Rules. *Workshop de Inteligencia Artificial. Jornadas Chilenas de Computación 2007*. Iquique, Chile. Noviembre de 2007
- [Ken95] Kenedy, Eberhart. Particle Swarm Optimization. *IEEE International Conference on Neural Networks*. Vol IV, pp.1942-1948. Australia 1995.
- [Lan00] Lanzarini. Reconocimiento de Patrones en Imágenes Médicas utilizando Redes Neuronales. *Journal of Computer Science and Technology*. Vol.4 . Dic 2000.
- [Lan04] Lanzarini, Yanivello. Reconocimiento de Comandos Gestuales utilizando GesRN. *V Workshop de Agentes y Sistemas Inteligentes*. CACIC 2004. Bs.As. Argentina. 2004. ISBN 987-9495-58-6
- [Lez07] Leza, Lanzarini. Aprendizaje de Juegos mediante cúmulos de partículas. *VIII Workshop de Agentes y Sistemas Inteligentes*. CACIC 2007. Corrientes, Argentina. 2007.
- [Oli05] Olivera y Lanzarini. Cyclic Evolution. A new strategy for improving controllers obtained by layered evolution. *Journal of Computer Science and Technology*. Vol 4,

nro. 1. 2005. Pags 211-217.

- [Ose06] Osella Massa, Vinuesa, Lanzarini. Modular Creation of Neuronal Networks for Autonomous Robot Control. *Ibero-American Journal of Artificial Intelligence*. Vol. 11, No. 35 (2007), pp. 43-53.
- [Shi07] Shichang, Ajito, Guiyong, Hongbo. A Particle Swarm Optimization Algorithm for Neighbor Selection in Peer-to-Peer Networks. *6th International Conference on Computer Information. Systems and Industrial Management Applications (CISIM'07)*. Pp. 166-172. ISBN 0-7695-2894-5. June 2007.
- [Shi99] Shi Y., Eberhart R. An empirical study of particle swarm optimization. *IEEE Congress Evolutionary Computation*. pp.1945-1949. Washington DC, 1999.
- [Vin07a] Vinuesa, Osella Massa, Corbalán, Lanzarini. Continuous Evolution of Neural Modules of Autonomous Robot Controllers. *Jornadas Chilenas de Computación (JCC 2007)*. Iquique. Chile. Noviembre 2007.
- [Vin07b] Vinuesa, Lanzarini. Neural Networks Elitist Evolution. *29th Internacional Conference Information Technology Interfaces (ITI 2007)*. Dubrovnik. Croatia. 2007.